

ОПЫТ ОЦИФРОВКИ И ПУБЛИКАЦИИ ЧЕРЕЗ ПОРТАЛ GBIF ГЕРБАРИЯ QAR

Н.В. Иванова¹, М.Ю. Ишмуратова², М.П. Шашков^{1,2}, Е.В. Гаврилькова², Р.Т. Мусина², С.С. Тыржанова², А.Н. Мадиева², Т.М. Рымжан², С.У. Тлеукенова²

¹*Институт математических проблем биологии РАН - филиал ИПМ им. М.В. Келдыша РАН
Россия, 142290, г. Пущино, ул. Проф. Виткевича 1*

²*Карагандинский университет им. академика Е.А.Букетова
Республика Казахстан, 100024, г. Караганда, ул. Университетская 28/3
e-mail: Natalya.dryomys@gmail.com*

На современном этапе необходимость оцифровки данных о биоразнообразии и предоставление открытого доступа к ним не вызывает сомнений. Мировой опыт показывает, что размещение цифровой копии научной биологической коллекции в открытом доступе через тематические репозитории повышает ее сохранность, надежность и научное использование. Казахстан пока остается “белым пятном” на цифровой карте глобального биоразнообразия, не смотря на колоссальный объем материалов, накопленных поколениями исследователей его природы. В докладе представлен рабочий процесс по оцифровке гербария Карагандинского университета им. Е.А. Букетова QAR.

На подготовительном этапе гербарий был зарегистрирован в международном реестре Index Herbariorum и получил акроним QAR; университет зарегистрирован в глобальном репозитории цифровых данных о биоразнообразии GBIF; на сервере университета установлено программное обеспечение для публикации данных GBIF Integrated Publishing Toolkit. Затем на основе адаптированных материалов международного курса BioDATA было проведено обучение сотрудников методам оцифровки, стандартизации данных гербария и их публикации через портал GBIF.

На первом этапе все образцы были систематизированы по семействам и баркодированы, после чего специалистами кафедры ботаники

выполнена проверка корректности определения. Затем выполняли оцифровку этикеток. Информацию с этикеток вносили вручную, полностью, на языке оригинала (казахском или русском). После внесения данных выполняли первичную проверку каждой записи. На следующем этапе выполняли географическую привязку образцов. Для образцов, не имеющих географических координат мест сборов, была выполнена ручная геопривязка методом “точка-радиус” по текстовым описаниям. Всего выполнено >450 ручных геопривязок. Кроме того, для каждого места сбора указывали по-возможности страну (в соответствии с границами ООН), область и административный район по актуальному состоянию. На заключительном этапе выполняли проверку и верификацию оцифрованных данных (data cleaning). Были исправлены технические ошибки. Имена коллекторов и даты сборов выверены согласно спискам студентов, обучавшихся на факультете, таксономические данные выверены согласно GBIF Backbone Taxonomy.

Итоговый набор данных доступен через портал GBIF (DOI: 10.15468/3asvbx) и включает 12548 записей. QAR - первый полностью оцифрованным гербарий, как в Казахстане, так и в странах Центральной Азии. В текущем году запланировано фотографирование и размещение на платформе GBIF изображений образцов в дополнение к имеющимся этикеточным данным.